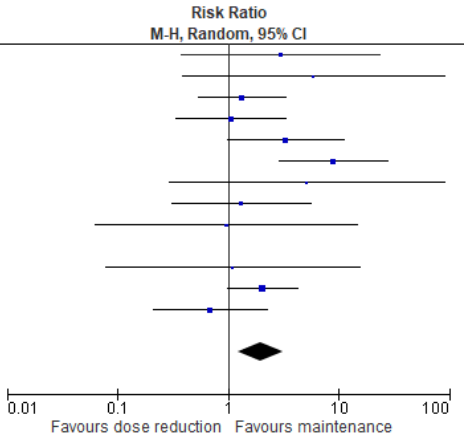
Supplementary Figure S4. Forest plot: Overall patients

1. Relapse/Hospitalization

1.1 Relapse (primary outcome)

Study or Subgroup	Dose reduction Events	Total	Maintenance Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Cookson 1987	3	9	1	9	4.4%	3.00 [0.38, 23.68]
Faraone 1989	8	22	0	7	2.7%	5.91 [0.38, 91.24]
Hogarty 1988	9	37	6	33	14.7%	1.34 [0.53, 3.36]
Inderbitzin 1994	5	20	4	17	11.2%	1.06 [0.34, 3.34]
Johnson 1987	9	28	3	31	10.4%	3.32 [1.00, 11.06]
Kane 1983	26	62	3	64	11.2%	8.95 [2.85, 28.05]
Newcomer 1992	3	14	0	10	2.5%	5.13 [0.29, 89.57]
Rouillon 2008	4	49	3	48	8.0%	1.31 [0.31, 5.53]
Takeuchi 2013	1	31	1	30	2.7%	0.97 [0.06, 14.78]
Uchida 2006	0	17	0	17		Not estimable
Volavka 2000	1	11	1	12	2.8%	1.09 [0.08, 15.41]
Wang 2010	19	120	10	129	18.7%	2.04 [0.99, 4.21]
Zhou 2018	4	37	6	38	10.7%	0.68 [0.21, 2.23]
Total (95% CI)		**457**		**445**	**100.0%**	**1.96 [1.23, 3.12]**
Total events	92		38			

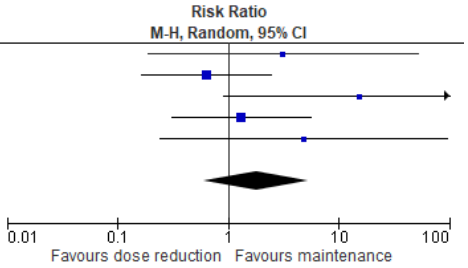
Heterogeneity: Tau² = 0.17; Chi² = 14.97, df = 11 (P = 0.18); I² = 27%
Test for overall effect: Z = 2.83 (P = 0.005)



1.2 Hospitalization

Study or Subgroup	Dose reduction Events	Total	Maintenance Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Faraone 1989	4	22	0	7	12.3%	3.13 [0.19, 51.93]
Inderbitzin 1994	3	20	4	17	33.4%	0.64 [0.17, 2.46]
Kane 1983	7	62	0	64	12.1%	15.48 [0.90, 265.33]
Rouillon 2008	4	49	3	48	31.1%	1.31 [0.31, 5.53]
Takeuchi 2013	2	31	0	30	11.1%	4.84 [0.24, 96.89]
Total (95% CI)		**184**		**166**	**100.0%**	**1.79 [0.60, 5.30]**
Total events	20		7			

Heterogeneity: Tau² = 0.45; Chi² = 5.69, df = 4 (P = 0.22); I² = 30%
Test for overall effect: Z = 1.04 (P = 0.30)

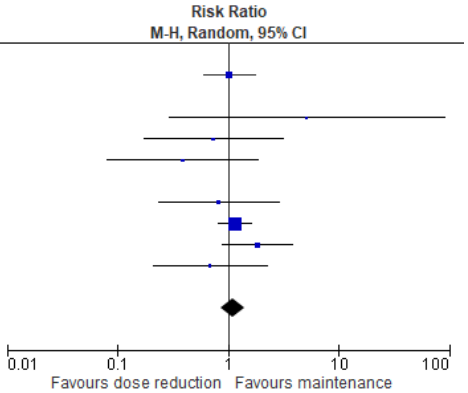


2. Study discontinuation

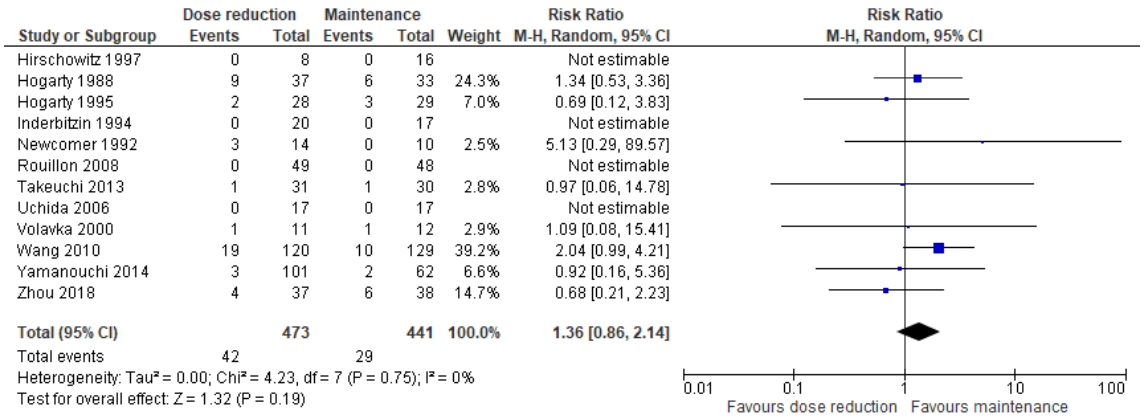
2.1 Study discontinuation due to all causes

Study or Subgroup	Dose reduction Events	Total	Maintenance Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Hirschowitz 1997	0	8	0	16		Not estimable
Hogarty 1988	16	37	14	33	21.2%	1.02 [0.59, 1.75]
Inderbitzin 1994	0	20	0	17		Not estimable
Newcomer 1992	3	14	0	10	0.8%	5.13 [0.29, 89.57]
Rouillon 2008	3	49	4	48	3.0%	0.73 [0.17, 3.11]
Takeuchi 2013	2	31	5	30	2.6%	0.39 [0.08, 1.84]
Uchida 2006	0	17	0	17		Not estimable
Volavka 2000	3	11	4	12	4.0%	0.82 [0.23, 2.87]
Wang 2010	44	120	41	129	52.4%	1.15 [0.82, 1.63]
Yamanouchi 2014	24	101	8	62	11.6%	1.84 [0.88, 3.84]
Zhou 2018	4	37	6	38	4.5%	0.68 [0.21, 2.23]
Total (95% CI)		**445**		**412**	**100.0%**	**1.11 [0.86, 1.42]**
Total events	99		82			

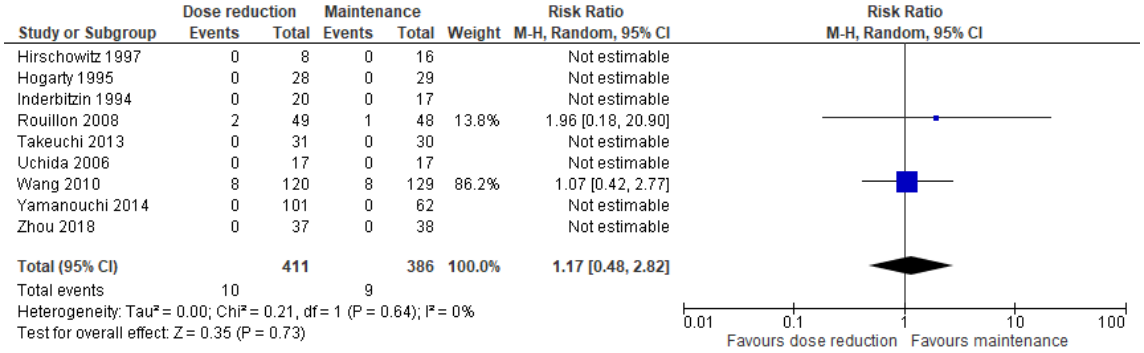
Heterogeneity: Tau² = 0.00; Chi² = 6.01, df = 7 (P = 0.54); I² = 0%
Test for overall effect: Z = 0.81 (P = 0.42)



2.2 Study discontinuation due to inefficacy

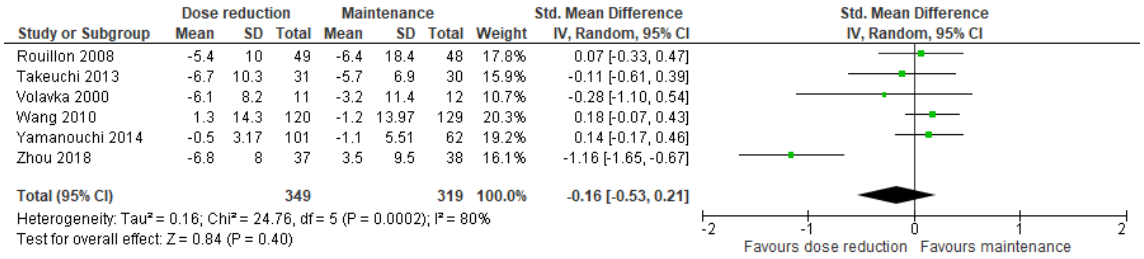


2.3 Study discontinuation due to intolerability

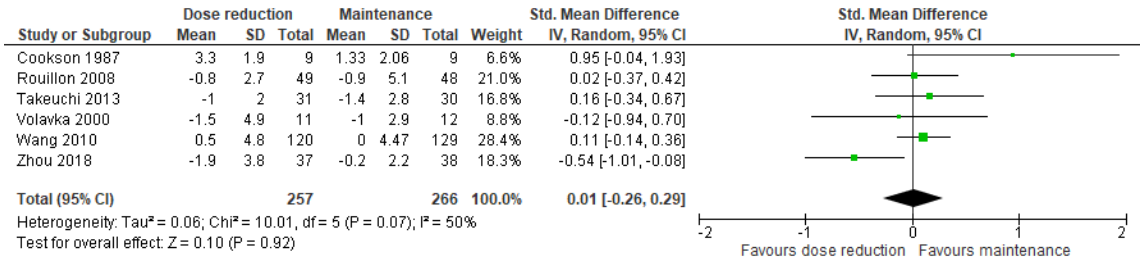


3. Psychopathology

3.1 Total



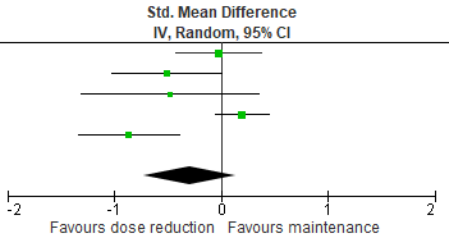
3.2 Positive symptoms



3.3 Negative symptoms

Study or Subgroup	Dose reduction Mean	SD	Total	Maintenance Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Rouillon 2008	-2	3.6	49	-1.9	5.6	48	21.9%	-0.02 [-0.42, 0.38]
Takeuchi 2013	-3	3.7	31	-1.3	2.8	30	19.6%	-0.51 [-1.02, 0.00]
Volavka 2000	-1.1	2.9	11	0.1	1.9	12	13.4%	-0.48 [-1.31, 0.36]
Wang 2010	-0.5	3.9	120	-1.3	4.19	129	24.8%	0.20 [-0.05, 0.45]
Zhou 2018	-2.4	7.4	37	3.7	6.6	38	20.3%	-0.86 [-1.34, -0.39]
Total (95% CI)			**248**			**257**	**100.0%**	**-0.29 [-0.72, 0.13]**

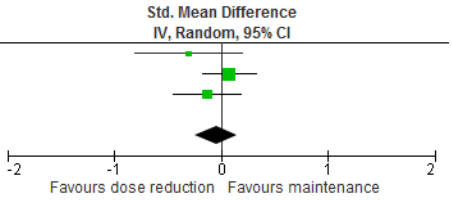
Heterogeneity: Tau² = 0.18; Chi² = 18.77, df = 4 (P = 0.0009); I² = 79%
Test for overall effect: Z = 1.35 (P = 0.18)



3.4 CGI-S

Study or Subgroup	Dose reduction Mean	SD	Total	Maintenance Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Takeuchi 2013	-0.2	0.7	31	0	0.6	30	14.3%	-0.30 [-0.81, 0.20]
Wang 2010	-0.1	1.1	120	-0.2	1.4	129	51.7%	0.08 [-0.17, 0.33]
Yamanouchi 2014	-0.1	0.55	101	0	1.06	62	34.0%	-0.13 [-0.44, 0.19]
Total (95% CI)			**252**			**221**	**100.0%**	**-0.05 [-0.24, 0.15]**

Heterogeneity: Tau² = 0.00; Chi² = 2.21, df = 2 (P = 0.33); I² = 9%
Test for overall effect: Z = 0.46 (P = 0.65)

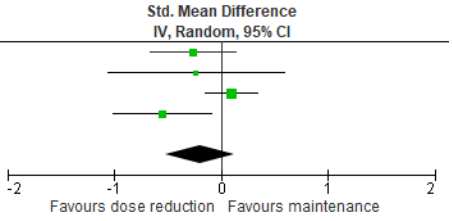


4. Adverse effects

4.1 SAS

Study or Subgroup	Dose reduction Mean	SD	Total	Maintenance Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Rouillon 2008	-0.2	1.4	49	0.3	2.3	48	27.4%	-0.26 [-0.66, 0.14]
Volavka 2000	-0.1	1.8	11	0.3	1.5	12	11.5%	-0.23 [-1.06, 0.59]
Wang 2010	-0.1	2.5	120	-0.3	1.7	129	37.1%	0.09 [-0.15, 0.34]
Zhou 2018	-1.4	3.5	37	0.2	2.1	38	24.0%	-0.55 [-1.01, -0.09]
Total (95% CI)			**217**			**227**	**100.0%**	**-0.20 [-0.52, 0.12]**

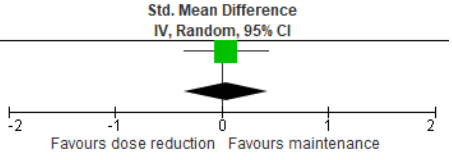
Heterogeneity: Tau² = 0.06; Chi² = 6.71, df = 3 (P = 0.08); I² = 55%
Test for overall effect: Z = 1.20 (P = 0.23)



4.2 BARS

Study or Subgroup	Dose reduction Mean	SD	Total	Maintenance Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Rouillon 2008	0.02	1.2	49	-0.04	1.6	48	100.0%	0.04 [-0.36, 0.44]
Total (95% CI)			**49**			**48**	**100.0%**	**0.04 [-0.36, 0.44]**

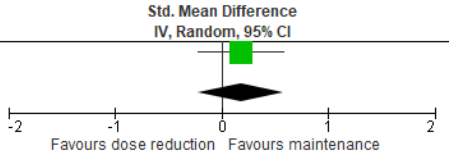
Heterogeneity: Not applicable
Test for overall effect: Z = 0.21 (P = 0.84)



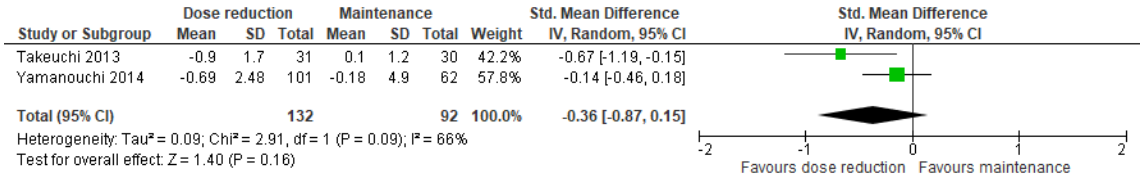
4.3 AIMS

Study or Subgroup	Dose reduction Mean	SD	Total	Maintenance Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Rouillon 2008	-0.1	1.6	49	-0.4	1.7	48	100.0%	0.18 [-0.22, 0.58]
Total (95% CI)			**49**			**48**	**100.0%**	**0.18 [-0.22, 0.58]**

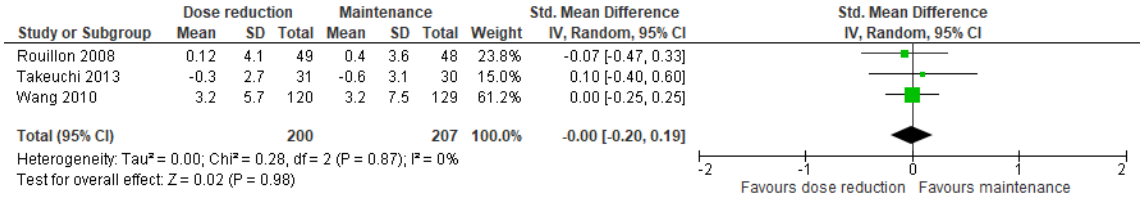
Heterogeneity: Not applicable
Test for overall effect: Z = 0.89 (P = 0.38)



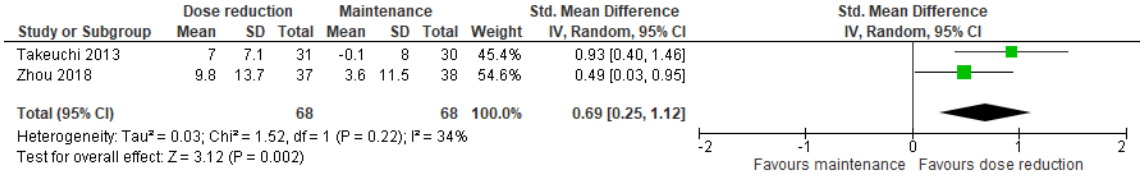
4.4 DIEPSS



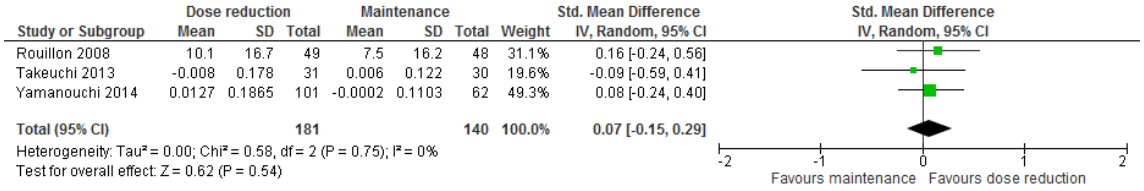
4.5 Body weight



4.6 Neurocognition



4.7 QOL



Abbreviations: AIMS, Abnormal Involuntary Movement Scale; BARS, Barnes Akathisia Rating Scale; CGI-S, Clinical Global Impressions – Severity scale; DIEPSS, Drug-Induced Extrapyramidal Symptoms Scale; QOL, quality of life; SAS, Simpson-Angus Scale